

Harish Kumar K

Full Stack & AI Engineer | AI-native Applications | Platform Engineering | 5+ Years Experience

📍 Vellore, India

📞 +91 8778257327

✉ harishkumar.vellore@gmail.com

🌐 harishkumar.info

📝 Blog

🌐 linkedin.com/in/harishkumarko25

🐙 github.com/thelonewolf123

Summary

AI Engineer with 5+ years building production-grade applications, real-time systems, and distributed infrastructure. Deep expertise in **LLM architecture, AI agent frameworks (LangGraph), observability, and systems engineering**. **67+ pages of technical documentation** and **12 published articles** on AI systems, real-time infrastructure, and distributed architecture.

Built four production-grade AI systems as personal projects: **Mail Pilot** (AI-native email with CopilotKit), **Quartermaster** (conversational AI agent with 26 tools), **Aegis** (AI-powered debugging with LangGraph), and **Stardust** (container platform on AWS Spot). Scaled a real-time platform to **7,000+ users** with paying customers. Delivered measurable impact: **150% booking growth, 15min → 5s latency reduction, 55% cost optimization**. Invited speaker at Chennai Web Conf (2025).

Experience

PickYourTrail

Software Developer (SDE II)

11/2024 - Present

Chennai, India

- Migrated React Native **0.76 → 0.82**, improving **performance and stability**.
- Revamped **hotel search and booking flow**, resulting in a **150% increase in bookings**.
- Built **cancellation & reschedule policy module** (core business workflow).
- Automated **GST invoice retrieval from Indigo flights (100-400 daily bookings)**.

BahiKhata Inc

Fullstack Developer

11/2023 - 11/2024

Noida, India

- Enabled **offline and multi-session sync** for MSME transaction recording across devices without internet.
- Developed **SMS parsing** to automatically detect and record transactions from bank messages.

Klenty

Full Stack Developer - Infra

02/2023 - 11/2023

Chennai, India

- Implemented **Zipkin distributed tracing** for system-wide observability.
- Built **real-time mail tracking system** → reduced latency from **15 min → 5 sec**.
- Designed **request ID attribution** for MongoDB queries enabling exact user-to-code-path tracing.
- Built **slow query detection pipeline** to identify and resolve performance bottlenecks.

Codedamn

Full Stack Developer - L3

08/2022 - 02/2023

Bangalore, India

- Built **DSA playground** for executing coding problems.
- Implemented **real-time collaborative coding mode** (multi-user editor).
- Built **autocomplete and IntelliSense system** for browser IDE.
- Optimized container allocation leading to a **55% cost reduction**.

Facilio

Member of Technical Staff

03/2021 - 07/2022

Chennai, India

- Built advanced **pivot table system** with conditional formatting and script columns for cross-module analytics.
- Designed **custom computation engine** enabling spreadsheet-like dynamic analysis.

Projects

Mail Pilot | AI-Native Email Client

[Live](#) [Docs](#)

- Built **AI-native email client** where users control everything through natural language.
- Designed **14 typed, Zod-validated AI tools** with a human-in-the-loop approval state machine for every email action.
- Built **file reasoning** (PDF, CSV, text) and **resume-to-job matching** powered by AI.
- Full technical documentation: **25+ docs across 9 categories** covering architecture, AI agent system, and security.

Quartermaster | Conversational AI Agent — 26 tools · 21 doc pages

[Live](#) [Docs](#)

- Built **conversational AI agent** running kirana store operations entirely through Telegram: billing, inventory, khata ledger, GST invoices, and analytics.
- Designed **26 typed tools** with LangGraph/deepagents, middleware pipeline, and subagent delegation.

- Implemented **multi-turn billing with GST computation**, automatic HSN/GST lookup, and customer credit management.
- Full technical documentation: **21 pages across 8 modules** covering architecture, modules, and deployment.

Aegis | **AI-Powered Production Debugging**

[Docs](#)

- Built **AI debugging system** that queries production databases, logs, and infrastructure via natural language.
- Multi-source correlation across PostgreSQL, CloudWatch, and IP geolocation with **real-time SSE streaming**.
- Implemented **5-layer defense-in-depth sandboxing**: Docker → iptables → NSJail → seccomp → read-only DB.
- Built **ComputeFlow**: agent writes JS on-the-fly for data processing instead of dumping rows into LLM context.

Stardust.app | **Container Deployment Platform — 6 microservices • 27 doc pages**

[GitHub](#) [Docs](#)

- Built **container provisioning engine on AWS Spot** with on-demand fallback and container checkpointing for disaster recovery.
- **Real-time logs via SSE**, WebSocket SSH, and auto Docker builds triggered by git push.
- IaC with Pulumi, **6 microservices** with RabbitMQ async orchestration, and 27 architecture docs.

WatchWithMe | **Real-time Watch Party Platform**

[Live](#)

- Built custom **P2P WebRTC library from scratch** (DHT routing, in-browser relays); migrated **4,000+ users** to LiveKit SFU with zero-downtime.
- Scaled to **7,000+ users** with paying customers. Increased signup completion **3x** through anonymous-first onboarding.
- Integrated voice + video chat and solved cross-platform sync, latency, and drift challenges.

Achievements

Invited Speaker | **Chennai Web Conf**

[Dec 2025](#)

- **Talk**: “Artemis: Personal AI OS” — an AI productivity system integrating task management, note-taking, reminders, Perplexity-powered research, and voice interaction.
- Presented to 100+ engineers on building AI-native applications and personal AI systems.

Technical Skills

Languages: TypeScript, JavaScript, Python

Frontend: React, React Native, Next.js, Tailwindcss, Shadcn

Backend: Node.js, Python, GraphQL, Express, FastAPI

Databases: PostgreSQL, MongoDB, Redis, MySQL

Gen AI & LLM: LangGraph, CopilotKit, OpenRouter, n8n, RAG Pipelines, Prompt Engineering, Vector Databases

Cloud & DevOps: AWS, Serverless, Docker, Pulumi

Core Strengths: Distributed Systems, Real-time Systems (WebSockets, Collaboration), Observability & Debugging, AI-powered Tools

Education

Thanthai Periyar Government Institute of Technology

B.E Electronics and Communication Engineering

[Vellore, India](#)

2017 – 2021